

## Classificação supervisionada com base em atributos cognitivos de cor para imagens de sensoriamento remoto

Daniel Capella Zanotta<sup>1</sup>  
Letícia Figueiredo Sartório<sup>1</sup>  
Fabiano Soares Dias<sup>2</sup>  
Bruna dos Santos Freitas<sup>1</sup>  
Tamires Pereira Pintado<sup>2</sup>

<sup>1</sup> Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul - IFRS  
Rua Eng. Alfredo Huch, 471, Rio Grande - RS, Brasil  
leticia.sartorio@outlook.com, [daniel.zanotta@riogrande.ifrs.edu.br](mailto:daniel.zanotta@riogrande.ifrs.edu.br)  
brunachristello@gmail.com

<sup>2</sup> Fundação Universidade Federal do Rio Grande - FURG  
Avenida Itália km 8, Rio Grande – RS – Brasil  
[fabianodias020@gmail.com](mailto:fabianodias020@gmail.com) tamires.pintado@riogrande.ifrs.edu.br

**Abstract.** This work aims at present a new classification method based on cognitive color attributes. Implemented in a supervised manner, the method assumes user collected samples for unlimited number of classes in a three dimensional space of attributes. Then, the samples are converted to HSV and plotted in a simplified two dimensional HSV diagram. The user has to select polygons on this reduced HSV space in order to generalize the scope of each class. Finally, the original image is all converted to HSV space and each element is considered in order to define if it lies in a region occupied by one class in the reduced HSV space. The main advantage of the proposed classification process is the power to emulate the human ability to acquiring knowledge and understanding through thought, experience, and the senses (cognitive aspects) used by photo interpreters during visual interpretation of remote sensing image targets. For this reason, the limited space of attributes is necessary, that is, to match the number of channels showed at screen. Experiments performed with an image marked by deforestation in Amazon were conducted by comparing the performance of several classification approaches. The method proved to be useful for noncomplex problems when simple approaches tend to show adequate results. Other advantages of the proposed method is its simplicity, its intuitiveness, besides its ability to generalize the sampling process in applications which collecting homogeneous samples is a difficult task.

**Palavras-chave:** remote sensing, image processing, color, classification, sensoriamento remoto, processamento de imagens, cor, classificação

### 1. Introdução

Abordagens alternativas como a análise de imagem baseada em objetos ou aplicação de modelos de mistura espectral tem desempenhado um papel importante no desenvolvimento de técnicas automáticas para classificação de imagens de sensoriamento remoto (Neubert e Meinel, 2003; Blashcke et al., 2008). Análises de mistura espectral podem, por exemplo, revelar proporções aproximadas de cada alvo de interesse a nível subpixel, conferindo objetividade aos valores radiométricos. Ao mesmo tempo, além das tradicionais características radiométricas, o agrupamento de pixels em objetos pode fornecer informações sobre textura, contexto e forma de determinados alvos que são cruciais para alcançar resultados adequados em algumas aplicações. Uma das vantagens do uso dessa abordagem é a proximidade entre os resultados gerados automaticamente e o trabalho produzido pela simples interpretação visual. Esta semelhança é induzida pela natureza humana de reconhecer objetos considerando atributos de cor, forma, textura e contexto (Navulur, 2006).

Mesmo considerando vantagens como rapidez e baixo custo apresentadas por análises automáticas de imagens, em algumas aplicações específicas, no entanto, classificações

manuais com base em interpretação visual podem apresentar resultados mais condizentes com as expectativas do estudo. Exemplos dessas aplicações podem ser encontrados em ações governamentais e mapeamentos oficiais que exigem alta precisão e confiabilidade dos resultados, como relatórios técnicos, alguns mapeamentos ambientais ou levantamentos urbanos de precisão. Nesse contexto, produtos derivados de técnicas automáticas ainda enfrentam certo descrédito por parte dos usuários finais ou tomadores de decisão quando se tratam de questões delicadas da administração. Pode-se citar, por exemplo, o cálculo do desflorestamento da Amazônia, estudos de impacto ambiental (EIA/RIMA), mapeamento para regularizações fundiárias, etc., onde a qualificação/quantificação precisa e a localização exata de áreas estudadas são informações críticas e essenciais.

As razões que impõem o uso cauteloso de produtos gerados por técnicas automáticas é a dificuldade enfrentada para adaptar regras de decisão a uma ampla gama de possibilidades, que podem, por sua vez, serem gerenciadas com facilidade pelo cérebro humano. Intérpretes visuais empregam sua experiência e treinamento para definir se um determinado conjunto de pixels pertence ou não a uma classe predefinida. O reconhecimento humano de alvos é realizado basicamente através de um processo de comparação com informações anteriores seguido de generalização (Hilbert, 2000). Esse processo é comprovado por experimentos envolvendo o treinamento com um conjunto limitado de amostras e posterior classificação do conjunto completo de dados [Shao e Lunetta, 2012]. As técnicas de classificação de imagens encontradas na literatura de sensoriamento remoto que seguem esse procedimento são conhecidas como supervisionadas. Esse processo geralmente é feito em domínios diferentes quando o usuário não pode visualizar todos os atributos disponíveis simultaneamente (por exemplo, o usuário seleciona amostras na tela e os dados são treinados e classificados usando dimensões adicionais). A generalização dos dados em um domínio mais completo do que aquele que as amostras foram coletadas pode produzir variações na definição das fronteiras de decisão entre classes. Essa instabilidade causada pela falta de correspondência entre o espaço amostral e o espaço de treinamento dos classificadores supervisionados produz efeitos como *overfitting* (excesso de ajuste do classificador aos dados amostrados), comprometendo o grau de generalização do processo de reconhecimento (Tso e Mather, 2009).

Neste trabalho, um processo de classificação automático baseado na cor dos alvos é proposto e testado em problemas que contam com atributos limitados de treinamento. Ao utilizar o mesmo espaço de atributos disponibilizado na tela para o usuário na coleta de amostras, o classificador proposto produz, em duas etapas, fronteiras de decisão baseadas em Matiz, Saturação e Intensidade dos alvos de interesse. O processo não impede que o usuário visualize a imagem em qualquer composição de falsa cor e colete as amostras das classes nessas condições. Dentre as vantagens do método proposto, destacam-se a possibilidade de trabalhar no mesmo espaço de atributos apresentado para visualização na tela e a maior correspondência visual dos dados classificados com a visualização da imagem na tela do computador. Além disso, a forma de implementação do método proposto torna a técnica insensível a ruídos apresentados por amostras heterogêneas ou mal coletadas. Mapeamentos que são tradicionalmente executados por pura interpretação visual, como detecção de desflorestamento, identificação de queimadas, e outros artefatos de tonalidade bem definida tendem a encontrar consonância com o método proposto, uma vez que o nível de complexidade do método é baixo e seu grau de generalização é alto.

## 2. Espaço RGB e HSV

Antes de apresentar o funcionamento da técnica de classificação proposta, é importante que os sistemas de cor RGB e HSV sejam caracterizados. Muito explorados em fotografia

convencional, mas pouco em aplicações de sensoriamento remoto, os sistemas de cor apresentam maneiras diferentes de representar uma determinada cor. No espaço RGB, as três únicas cores percebidas pelo olho humano, vermelho, verde e azul, são utilizadas para formar as outras cores existentes no espectro visível. Uma forma de representação clássica desse sistema é o diagrama de coordenadas cartesianas em três dimensões (Figura 1a). Já o sistema HSV define as cores disponíveis no espectro visível através de três variáveis que são Matiz, Saturação e Intensidade. A Matiz se referindo à tonalidade, a Saturação ao grau de pureza da cor, e a Intensidade à luminosidade. A representação do espaço HSV se dá, geralmente, através de um cone (Fig. 1b), sendo a Matiz uma medida angular  $[0,2\pi]$ , a Saturação uma medida radial  $[0,r]$  e a intensidade uma medida linear  $[0,v]$ .

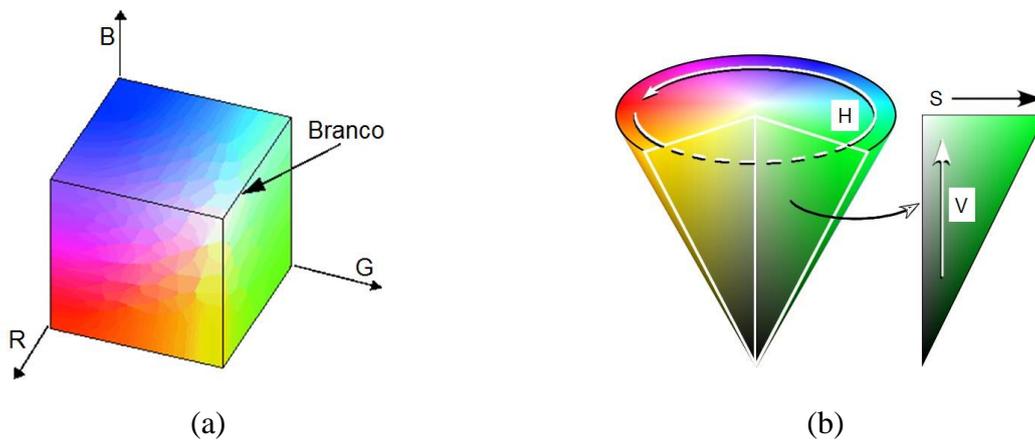
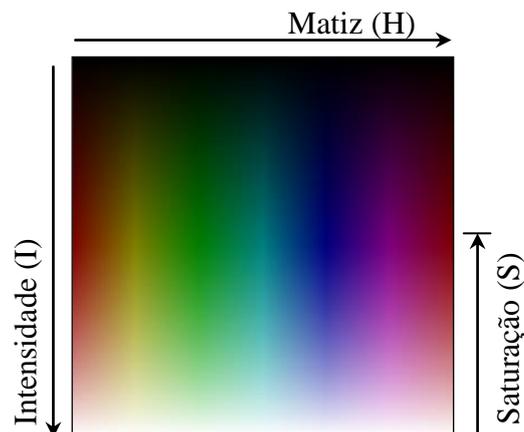


Figura. 1 Representação dos sistemas (a) RGB e (b) HSV.

Tanto o sistema RGB quanto o sistema HSV são melhores representados em espaços tridimensionais. No entanto, a plenitude das opções de cor oferecidas por tais figuras sólidas não é convenientemente visualizável, levando alguns autores a separar a definição de uma cor em duas etapas bidimensionais distintas. No caso do sistema HSV, matiz e saturação (H e S) são geralmente escolhidas primeiro, sendo a intensidade (V) da cor definida posteriormente. Naturalmente, variações de tonalidade e intensidade de cor são mais facilmente manipuláveis separadamente no espaço HSV do que no espaço RGB.

No presente estudo uma redução do espaço HSV tradicional foi realizada para simplificar o processo de visualização das cores disponíveis na tela do computador (Figura 2). Ainda assim, as variações de Matiz, Saturação e Intensidade foram preservadas, como pode ser visto em suas componentes apresentadas nas Figuras 2a, 2b e 2c, respectivamente.



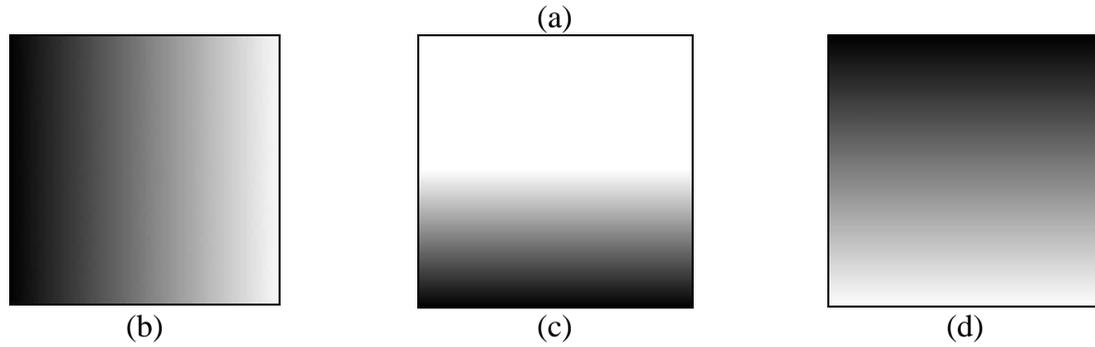


Figura. 2 Representação do sistema HSV reduzido (a) e suas componentes H (b), S (c) e V (d).

### 3. Técnica de classificação baseada em cor

Assim como em metodologias supervisionadas tradicionais, a primeira etapa da técnicas proposta é a coleta de amostras de cada uma das classes de interesse para o estudo. Na definição do conjunto amostral o usuário deve procurar seguir algumas recomendações básicas como escolher regiões representativas da classe, não coletar amostras muito diferentes para uma mesma classe, não deixar amostras de uma determinada classe de fora do problema, evitar regiões contendo elementos pertencentes a mais de uma classe, e coletar um número de elementos suficiente para garantir a correta caracterização de cada atributo considerado (Tso e Mather, 2009). Como será visto a seguir, alguns desses problemas de amostragem serão mitigados na abordagem proposta nesse estudo. A conversão entre os espaços RGB e HSV pode ser realizada a partir das seguintes equações de suporte.

$$\begin{aligned}
 R^{\wedge} &= R/255 & C \max &= \max(R^{\wedge}, G^{\wedge}, B^{\wedge}) \\
 G^{\wedge} &= G/255 & C \min &= \min(R^{\wedge}, G^{\wedge}, B^{\wedge}) \\
 B^{\wedge} &= B/255 & \Delta &= C \max - C \min
 \end{aligned} \tag{1}$$

Onde  $R$ ,  $G$  e  $B$  são os valores das bandas selecionadas para os canais do vermelho, verde e azul, respectivamente. Os valores de  $R^{\wedge}$ ,  $G^{\wedge}$  e  $B^{\wedge}$  correspondem à normalização dos canais originais.  $C \max$  e  $C \min$  referem-se ao maior e menor valor registrado para cada pixel e  $\Delta$  corresponde à diferença entre os dois valores. A determinação de  $H$ ,  $S$  e  $V$  pode ser realizada a partir das seguintes equações.

$$H = \begin{cases} 0^{\circ} & \Delta = 0 \\ 60^{\circ} \times \left( \frac{G^{\wedge} - R^{\wedge}}{\Delta} \bmod 6 \right) & , C \max = R^{\wedge} \\ 60^{\circ} \times \left( \frac{B^{\wedge} - R^{\wedge}}{\Delta} + 2 \right) & , C \max = G^{\wedge} \\ 60^{\circ} \times \left( \frac{R^{\wedge} - G^{\wedge}}{\Delta} + 4 \right) & , C \max = B^{\wedge} \end{cases} \tag{2}$$

$$S = \begin{cases} 0 & , C \max = 0 \\ \frac{\Delta}{C \max} & , C \max \neq 0 \end{cases}$$

$$V = C \max$$

A metodologia proposta subentende a conversão da imagem de entrada de RGB para HSV nas 3 bandas escolhidas, não necessariamente aquelas correspondentes ao visível do espectro eletromagnético. O treinamento do classificador inclui duas etapas distintas, denominadas (1) Processo de Amostragem e (2) Generalização do Problema.

### 3.1 Processo de amostragem

O procedimento inicial consiste na coleta de mostras para cada uma das classes envolvidas no problema a fim de produzir uma representação inicial de cada conjunto amostral no espaço HSV reduzido. A Figura 3 apresenta uma simulação do que seriam amostras de três classes diferentes e seu respectivo posicionamento no espaço HSV reduzido. Nessa representação, uma figura do satélite CBERS-4 (Figura 3a) foi utilizada para extrair algumas tonalidades de cores específicas como dourado, verde e lilás. As correspondentes matiz, saturação e intensidade de cada pixel das amostras foram representadas no diagrama HSV da Figura 3b, sendo cada amostra representada por uma cor diferente, escolhida de maneira intuitiva.

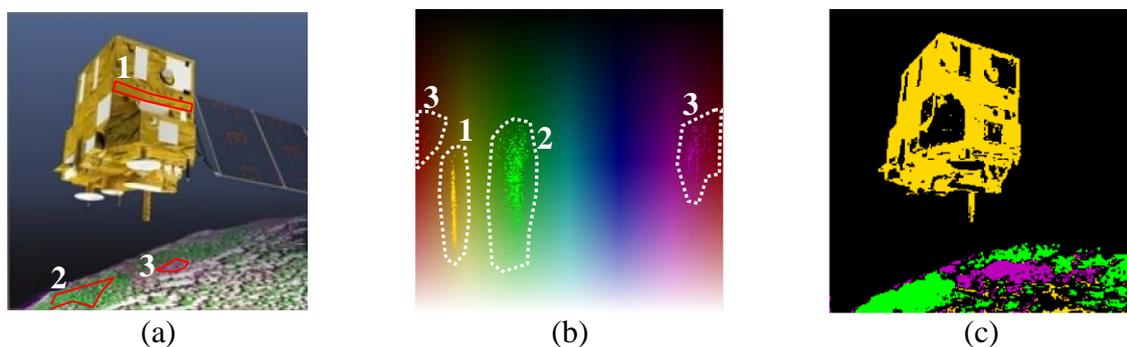


Figure 3. Processo de classificação supervisionada. (a) coleta de amostras de diferentes alvos contidos na cena, baseado em tonalidade de cor, (b) representação dos pixels amostrados no espaço de HSV reduzido, (c) classificação final através do método baseado em cor proposto.

### 3.2 Generalização do problema

A aplicação da metodologia prossegue com a seleção de polígonos aumentando a abrangência de cada classe diretamente no diagrama HSV (Fig. 3b). Essa etapa objetiva aumentar a generalização do classificador, contemplando não apenas os tons representados por pontos amostrados pelo usuário, mas também colorações vizinhas, semelhantes a estas. A seguir, a imagem inteira é convertida para HSV e sua análise se dá de acordo com o diagrama construído pelo usuário (Fig. 3b). A regra de decisão é simples: se algum pixel, em sua representação HSV, apresentar valores que correspondam a regiões selecionadas, o pixel é atribuído à classe da respectiva região. Como pode ser notado, o método de classificação proposto não realiza uma análise exaustiva, sendo possível que alguns elementos não sejam atribuídos a nenhuma das classes selecionadas pelo usuário. Nesse caso, o pixel recebe o valor zero na classificação (preto na imagem da Fig. 3c). A fim de aumentar a quantidade de pixels classificados, o usuário deve (1) aumentar o número de classes do problema, ou (2) aumentar o tamanho das regiões de abrangência na segunda etapa do treinamento.

### 3. Experimentos e Análise

A fim de verificar o desempenho do método de classificação proposto, um experimento com uma imagem *TM-Landsat-5* cobrindo uma região de Rondônia foi realizado. O *subset* da imagem (Fig. 4a) contém basicamente três classes: floresta, solo exposto e vegetação rasa. As duas últimas classes correspondem a sucessões de desflorestamento. Para fins de comparação, o método proposto foi testado com métodos tradicionais (máxima verossimilhança gaussiana e mínima distância euclidiana) e métodos não paramétricos (árvore de decisão e *Support Vector Machines*).

A área escolhida para o experimento corresponde a uma região da Amazônia brasileira contendo fragmentos de desflorestamento, representado por solo exposto (1), vegetação rasa (2) e floresta primária (3). Amostras dessas classes foram coletadas de maneira a representar cada uma delas por polígonos representados pela Figura 4.

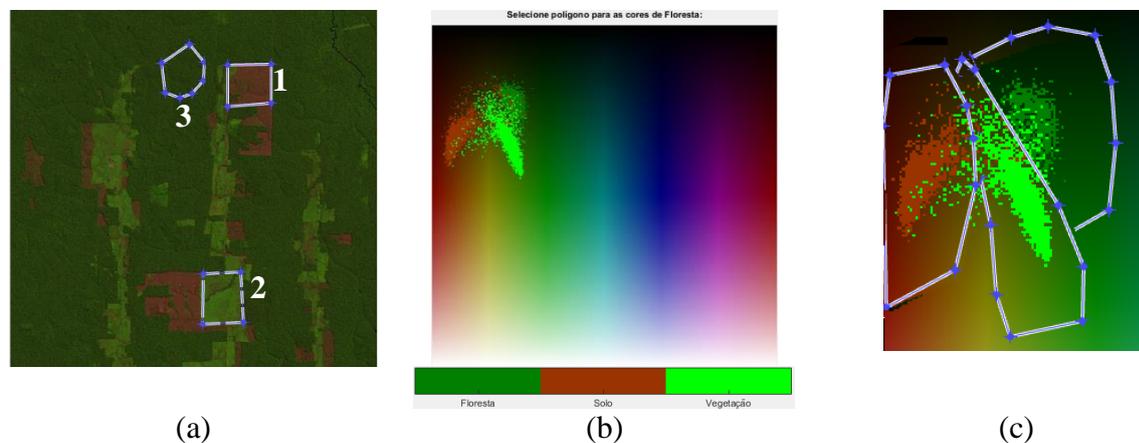


Figura 4. (a) Seleção de amostras de treinamento para floresta (3), solo exposto (1) e vegetação rasa (2) em imagem *TM-Landsat-5*, 5(R)4(G)3(B); (b) plotagem no espaço de atributos dos elementos amostrais convertidos para HSV para cada classe; (c) detalhe do espaço HSV reduzido mostrando delimitação das fronteiras de decisão na etapa de generalização do método. Os polígonos representados nas figuras são produzidos pelo usuário.

Após a etapa de treinamento, a imagem foi classificada a partir dos classificadores escolhidos para os testes. As classificações foram avaliadas com medidas de acurácia global apresentando: método baseado em cor 94,5%, Máxima verossimilhança gaussiana 93,8%, Mínima distância Euclidiana 93,3%, Árvore de Decisão 95,7% e SVM 96,1%. Os mapas de classificação produzidos são mostrados na Figura 5. Dada a simplicidade do método baseado em cor proposto, é de extrema valia que o resultado apresentado por essa abordagem tenha alcançado uma acurácia global de 94,5%, uma vez que a detecção foi realizada com um conjunto limitado de atributos quando comparado com os demais classificadores. Mesmo frente a métodos complexos e modernos como SVM e Árvore de Decisão, o método não apresenta resultados tão distantes. Vale observar o potencial do método proposto, em sua segunda etapa do treinamento (generalização), a possibilidade de o usuário facilmente corrigir possíveis misturas entre classes dentro de uma mesma amostra.

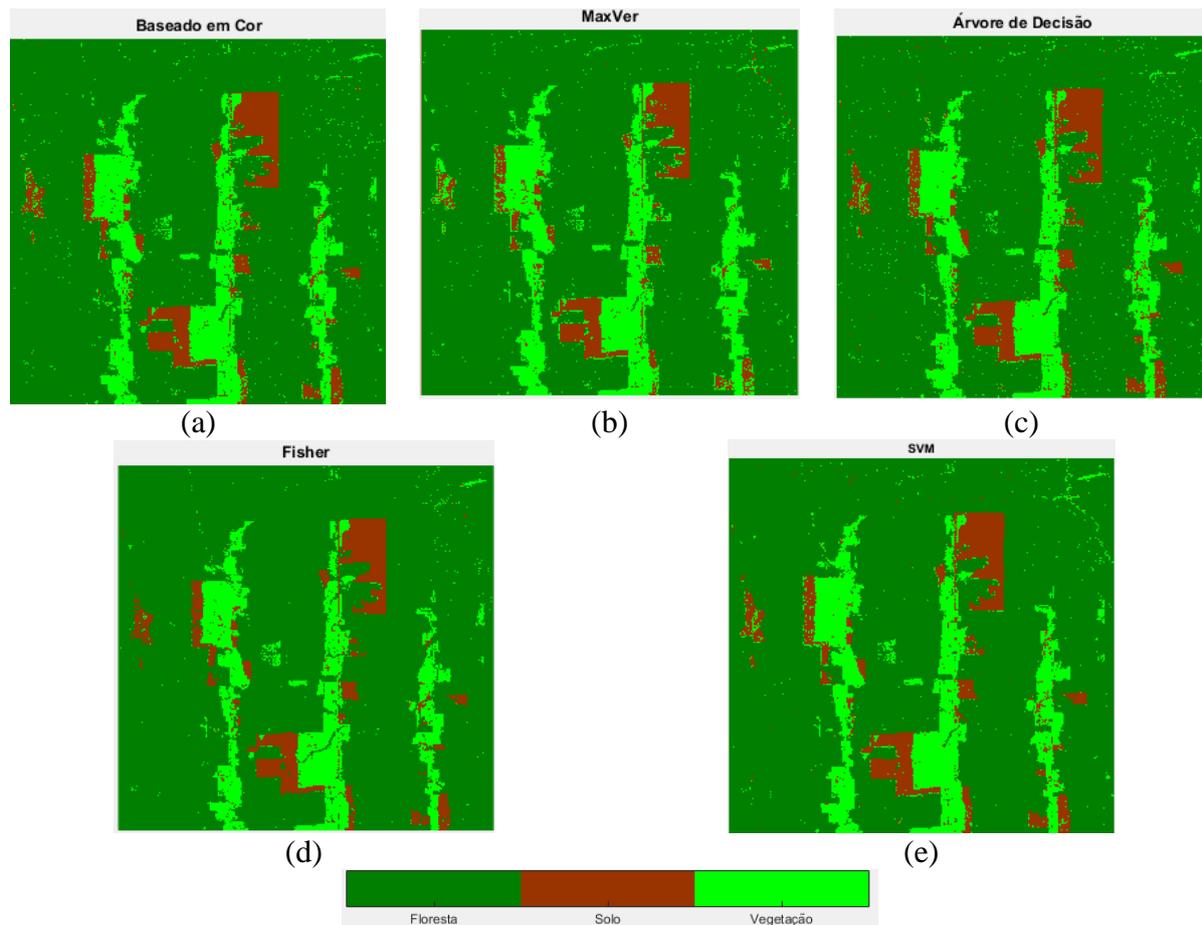


Figura 5. Mapas de classificação para a área de estudo com comparações do método proposto (a) com Máxima verossimilhança gaussiana (b), Árvore de Decisão (c), Mínima distância Euclidiana (d), e SVM (e).

#### 4. Conclusões

Neste trabalho um método de classificação supervisionado baseado em atributos de cor é proposto e testado para uma cena contendo regiões de desmatamento na Amazônia brasileira. Classificadores tradicionais exploram um grande conjunto de atributos em seu treinamento, o que acaba impedindo que o usuário tenha um completo entendimento do grau com que cada atributo influencia a classificação dos alvos na cena. A abordagem proposta é inspirada na classificação por interpretação visual realizada por foto-interpretas em alguns tipos de mapeamento específicos, onde estes visualizam os alvos em tela com uma determinada composição colorida (três canais visíveis) e delimitam os alvos de interesse com base em suas cores características. O método supervisionado proposto subentende a coleta de amostras de cada classe em um conjunto limitado de atributos (e.g. bandas espectrais), apenas três. Esses atributos são convertidos para o espaço HSV reduzido, independente de representarem ou não cores verdadeiras na visualização. A partir desse ponto, o usuário pode selecionar as regiões de abrangência de cada classe a fim de definir as fronteiras de decisão nesse mesmo espaço. Os resultados do experimento mostraram que as acurácias globais são superiores quando comparadas com métodos simples, mas inferiores quando comparadas com métodos mais modernos de classificação. A maior vantagem do método é a capacidade de generalização do problema, fruto da realização do processo de treinamento em duas etapas distintas: amostragem e generalização. Assim, amostras não homogêneas contendo mais de uma classe podem ser corrigidas na segunda etapa do processo. Se projeta que a maior aplicação do

método seriam aqueles problemas que são comumente resolvidos através da classificação por interpretação visual, como detecção de desmatamento, queimadas, mapeamento urbano, etc., onde o método tende a encontrar maior consonância.

## Referências

- Neubert, M., Meinel, G. **Evaluation of segmentation programs for high resolution remote sensing applications.** *International ISPRS Workshop High resolution mapping from space*, 2003.
- Blaschke, T., Stefan, L., Geoffrey, H. **Object-based image analysis: spatial concepts for knowledge-driven remote sensing applications.** *Springer Science & Business Media*, 2008.
- Navulur, K. Multispectral image analysis using the object-oriented paradigm. CRC press, 2006.
- Hilbert, C. J. **Protótipo de software para auxiliar no aprendizado das cores e formas geométricas.** *Trabalho de Conclusão de Curso (Bacharelado em Ciências da Computação). Centro de Ciências Exatas e Naturais. Universidade Regional de Blumenau.* 2000.
- Shao, Y., & Lunetta, R. S. (2012). **Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points.** *ISPRS Journal of Photogrammetry and Remote Sensing*, 70, 78-87.
- Mather, Paul; Tso, Brandt. **Classification Methods for Remotely Sensed Data.** 2. ed. London: CRC Press, 2009.